

AD\_\_\_\_\_

AWARD NUMBER: W81XWH-12-1-0293

TITLE: High-Throughput Functional Validation of Progression Drivers in Lung  
Adenocarcinoma

PRINCIPAL INVESTIGATOR: Kenneth L. Scott, Ph.D.

CONTRACTING ORGANIZATION: Baylor College of Medicine  
Houston, TX 77030-3498

REPORT DATE: September 2013

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command  
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;  
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

| REPORT DOCUMENTATION PAGE  |                  |                          |                                      | Form Approved<br>OMB No. 0704-0188                     |  |
|--|------------------|--------------------------|--------------------------------------|--|--|
| Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. <b>PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.</b>  |                  |                          |                                      |  |  |
| 1. REPORT DATE<br>September 2013   |                  | 2. REPORT TYPE<br>Annual |                                      | 3. DATES COVERED<br>1 October 2012 – 30 September 2013 |  |
| 4. TITLE AND SUBTITLE<br>High-Throughput Functional Validation of Progression Drivers in Lung Adenocarcinoma   |                  |                          |                                      | 5a. CONTRACT NUMBER<br>W81XWH-12-1-0293                |  |
|  |                  |                          |                                      | 5b. GRANT NUMBER<br>W81XWH-12-1-0293                   |  |
|  |                  |                          |                                      | 5c. PROGRAM ELEMENT NUMBER                             |  |
| 6. AUTHOR(S)<br>Kenneth Scott, Ph.D.<br>Don Gibbons, M.D, Ph.D.<br><br>E-Mail: kls1@bcm.edu  |                  |                          |                                      | 5d. PROJECT NUMBER                                     |  |
|  |                  |                          |                                      | 5e. TASK NUMBER  |  |
|  |                  |                          |                                      | 5f. WORK UNIT NUMBER                                   |  |
| 7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)<br>Baylor College of Medicine<br>One Baylor Plaza<br>Houston, TX 77030-3498   |                  |                          |                                      | 8. PERFORMING ORGANIZATION REPORT NUMBER               |  |
| 9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)<br>U.S. Army Medical Research and Materiel Command<br>Fort Detrick, Maryland 21702-5012  |                  |                          |                                      | 10. SPONSOR/MONITOR'S ACRONYM(S)                       |  |
|  |                  |                          |                                      | 11. SPONSOR/MONITOR'S REPORT NUMBER(S)                 |  |
| 12. DISTRIBUTION / AVAILABILITY STATEMENT<br>Approved for Public Release; Distribution Unlimited   |                  |                          |                                      |  |  |
| 13. SUPPLEMENTARY NOTES  |                  |                          |                                      |  |  |
| 14. ABSTRACT<br>The primary objective of our <i>Early Investigator Synergistic Idea Award</i> is to establish a driver prioritization pipeline to functionally evaluate lung cancer genomics data to identify somatic driver aberrations, which beyond the handful of well-characterized genes like oncogenic <i>KRAS</i> , contribute to lung cancer progression. As outlined in our Annual Report, we have successfully constructed the necessary gene libraries for the screens outlined in our proposal (Aim 1), and we have optimized and initiated the proposed <i>in vitro</i> and <i>in vivo</i> screens (Aim 2). We have already identified several robust drivers of cell invasion, and we are on task to complete all screens in year two of this project as planned. Finally, we are putting into place the necessary tools for mechanistic studies of lead candidate genes (Aim 3) that includes our development of a novel isogenic human bronchial epithelial cell system that permits regulatable expression of the <i>KRAS</i> oncogene. We have made several technical improvements to our overall work plan and view this first year as a highly successful start to this project and further screening of lung cancer genomics data that will extend past the life of this particular award. |                  |                          |                                      |  |  |
| 15. SUBJECT TERMS<br>Lung cancer driver, oncogenes, high-throughput screens  |                  |                          |                                      |  |  |
| 16. SECURITY CLASSIFICATION OF:  |                  |                          | 17. LIMITATION OF ABSTRACT<br><br>UU | 18. NUMBER OF PAGES<br><br>12                          | 19a. NAME OF RESPONSIBLE PERSON<br>USAMRMC |
| a. REPORT<br>U   | b. ABSTRACT<br>U | c. THIS PAGE<br>U        |                                      |  | 19b. TELEPHONE NUMBER (include area code)  |

TABLE OF CONTENTS

|                                   | <u>Page</u> |
|-----------------------------------|-------------|
| Introduction.....                 | 2           |
| Body.....                         | 2           |
| Key Research Accomplishments..... | 8           |
| Reportable Outcomes.....          | 9           |
| Conclusion.....                   | 9           |
| References.....                   | 9           |
| Appendices.....                   | 10          |

## **INTRODUCTION**

Cancer cells are endowed with diverse biological capabilities driven by an ensemble of inherited, somatic and epigenetic aberrations. As we enter the era of personalized medicine, characterization of the cancer genome has begun and will continue to influence diagnostic and therapeutic decisions in the clinic. Genome profiling technologies are generating a compendium of genomic aberrations in major cancer lineages with the goal of identifying the most promising therapeutic targets and diagnostic biomarkers. The game changing output from these technologies is radically transforming cancer science. At the same time, these efforts are revealing the complexity of cancer genomes, which are comprised of causal “driver” aberrations and many more biologically neutral “passengers” that arise through the unstable nature of tumor genomes. While most cancers acquire one or more well-studied, high frequency driver events (e.g., mutations/gene copy number changes in *KRAS*, *TP53*, *EGFR*, *MYC*, *BRAF*, etc.), much less is known about the overly abundant low frequency (<5%) aberrations contributing to tumor progression and response to therapeutics. Comprehensive biological assessment of low frequency aberrations is difficult given their large number and the fact that they may either directly drive tumor progression or indirectly influence tumor behavior through modifying activities of other drivers like *KRAS*. Moreover, distinguishing driver events from passengers is further complicated by the fact that driver events are shaped by the specific biological context of a given cancer, including its tissue type, microenvironment and other host determinants including the immune system. The primary objective of our *Early Investigator Synergistic Idea Award* is to establish a **driver prioritization pipeline** to functionally evaluate lung cancer genomics data to identify somatic driver aberrations, which beyond the handful of well-characterized genes like oncogenic *KRAS* and *EGFR*, likely contribute to lung cancer progression, invasion, and metastasis. Our study integrates genetically engineered mouse models of lung cancer, genomics data generated by The Cancer Genome Atlas (TCGA) and functional screens to identify drivers of lung cancer progression. We are using these tools to implement a novel, scalable screening infrastructure that permits high-content, gain-of-function screens to accelerate validation of functional somatic aberrations. This work is possible through advances made in our laboratory that include (1) high-throughput, highly accurate modeling of somatic aberrations into our collection of over 32,000 sequence-verified human genes and (2) a novel molecular barcoding approach that facilitates cost-effective detection of driver events following *in vitro* and *in vivo* functional screens. Our Specific Aims are as follows: (1) Construction of a lung cancer somatic driver library; (2) Functional screens for drivers of lung cancer metastasis; (3) Clinicopathological prioritization and validation of top candidates. **Herein we describe our progress over the first 12 months of this project.**

## **BODY**

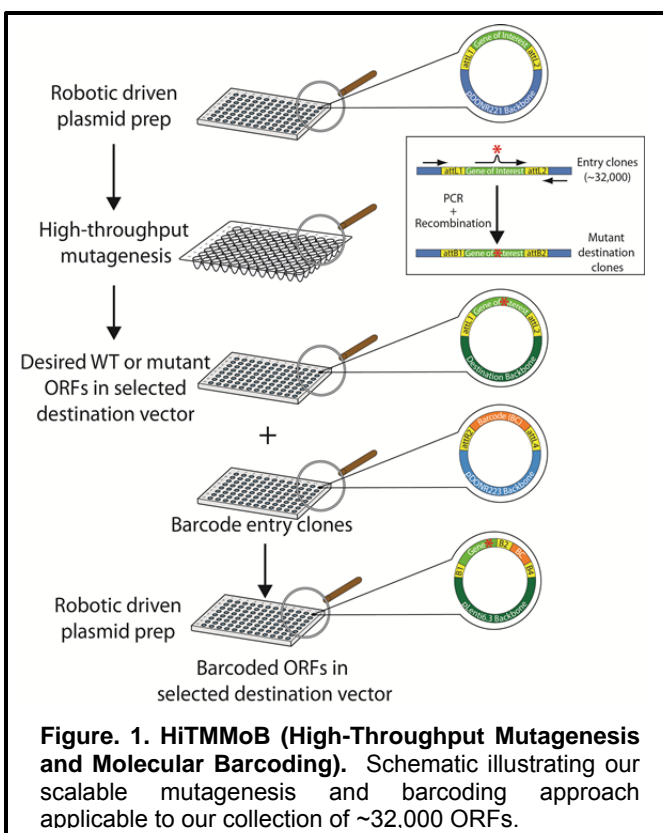
Based on our approved Statement of Work, our overall goals, timeline and progress for the first 12 months of this project were as follows:

### **Specific Aim 1: Construction of a lung cancer somatic driver library (proposed completion, months 3-6).**

As described in our application, we collaborated with Dr. Chad Creighton at Baylor College of Medicine to select a high-priority list of genes for functional screening. This list of 260 candidates was based upon available data from targeted sequencing of human lung adenocarcinomas and integration of those datasets with gene expression profiling data derived from tumors collected from our mouse models of lung cancer. The first task of this project was to create gene “libraries” comprised of those 260 candidates for our proposed *in vitro* and *in vivo* gain-of-function (GOF) screens. We were able to efficiently and entirely complete this first important Aim, due in part to the fact that our clone processing ability has significantly increased over the duration of this project by (1) increasing our wild-type gene collection from 19,000 to 32,000 open reading frame (ORF) clones and (2) streamlining workflow of our aberration construction pipeline, which we named High-Throughput

Mutagenesis and Molecular Barcoding (HiTMMoB; Fig.1). We have therefore been able to broaden the spectrum and number of aberrations chosen for our studies. We successfully generated an arrayed ORF library consisting of 279 sequence-verified wild-type and mutant ORF donor clones in 96-well plate format. These ORFs were subsequently recombined into the pLentiEF6.3-puro lentiviral vector specially constructed for this project. This vector is compatible with our flexible molecular barcoding technology (Fig. 1), thus each ORF is uniquely tagged with a 24-nucleotide DNA “barcode” followed by high-throughput DNA sequencing from bacteria to confirm each barcode and the identity of its associated ORF. Validated barcoded ORFs were re-arrayed into a 96-well format compatible with lentivirus production for the proposed studies.

Given the increase in our workflow capacity and the robustness of assay readouts (described below), we are currently working with Dr. Creighton to analyze the emerging data from TCGA project on lung adenocarcinoma to expand our set of genes for functional screening. To maximize the discovery potential in an area of critical clinical need, we are focusing on the co-incident mutations, copy number alterations, and gene expression changes unique to the mutant *Kras* population.



## **Specific Aim 2: Functional screens for drivers of lung cancer metastasis.**

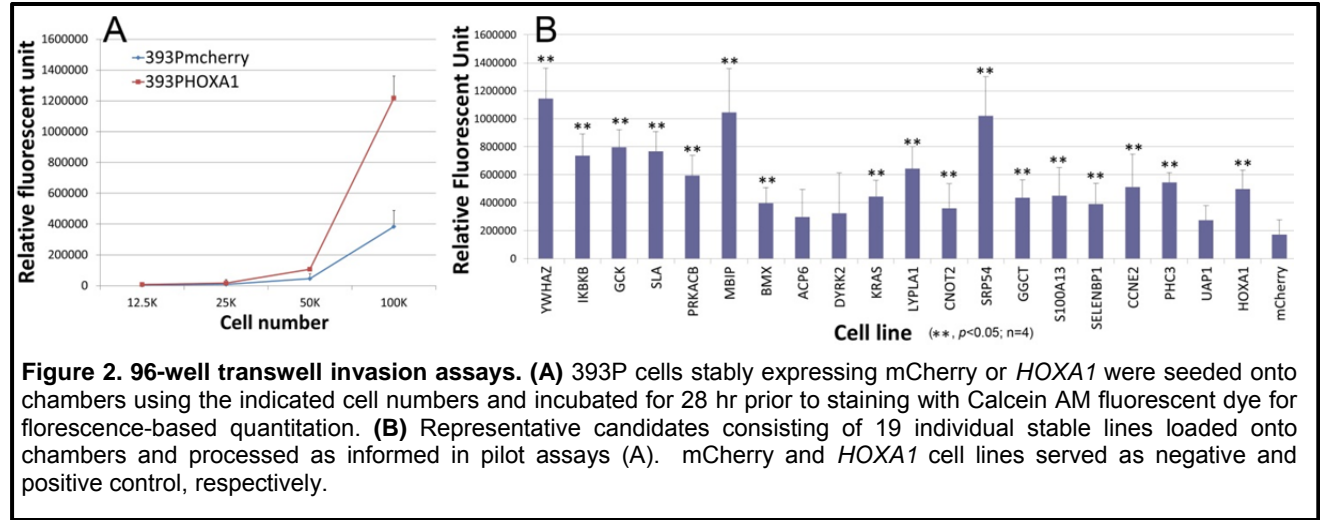
### **Subaim 2.1 *In vitro* screens for cell invasion and anoikis resistance (proposed completion, months 6-12)**

In this Subaim we propose to perform parallel *in vitro* screens for gene drivers of cell invasion and anoikis resistance using a 96-well format for cells transduced with single candidate genes from the ORF library constructed in Aim 1. We began by performing extensive optimization of all steps required for these screens, including establishing cell plating and screening conditions (pilot assays) for three separate cell lines: 393P and 393LN murine cell lines (described in our application) and a new cell line derived from human bronchial epithelial cells (HBECs; see description in Aim 3 below).

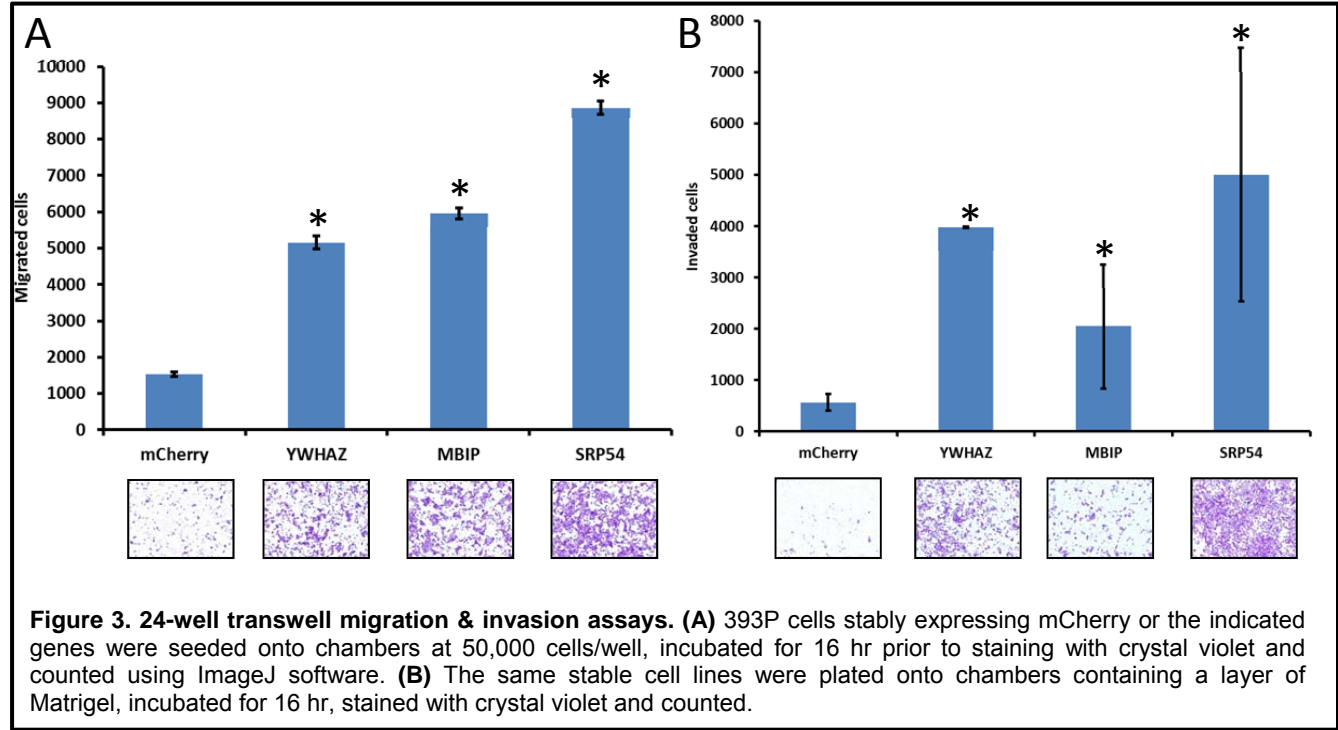
**Cell Invasion:** We first conducted a series of pilot assays employing a 96-well cell invasion assay used and described by us previously [1]. We assessed performance of the 393P, 393LN and HBEC cell lines stably transduced with negative (mCherry) and positive (*HOXA1*; [1],[2]) control drivers of cell invasion. These initial studies led us to choose 393P cells as the primary screening model for the invasion screens. This decision was based on the observations that 393P cells (1) form primary tumors upon inoculation of immunocompetent syngeneic animals but do not metastasize, thus permitting us to screen for metastasis drivers, (2) have low baseline invasion activity (Fig. 2A), and (3) can be stimulated to invade in the presence of a *bona fide* invasion driver, *HOXA1* (Fig. 2A).

Based on the successful pilot studies described above, we have initiated screening of our ORF library in 393P. Briefly, 393P cells are infected with lentivirus carrying the indicated genes, selected in puromycin for stable expression, and then screened for *in vitro* invasion using the established conditions (Fig. 2) and controls. Importantly, all stable cell lines are cryo-banked for (1) future use in functional and biochemical validation assays and (2) *in vivo* screening assays proposed in Subaim 2.2 and as described below.

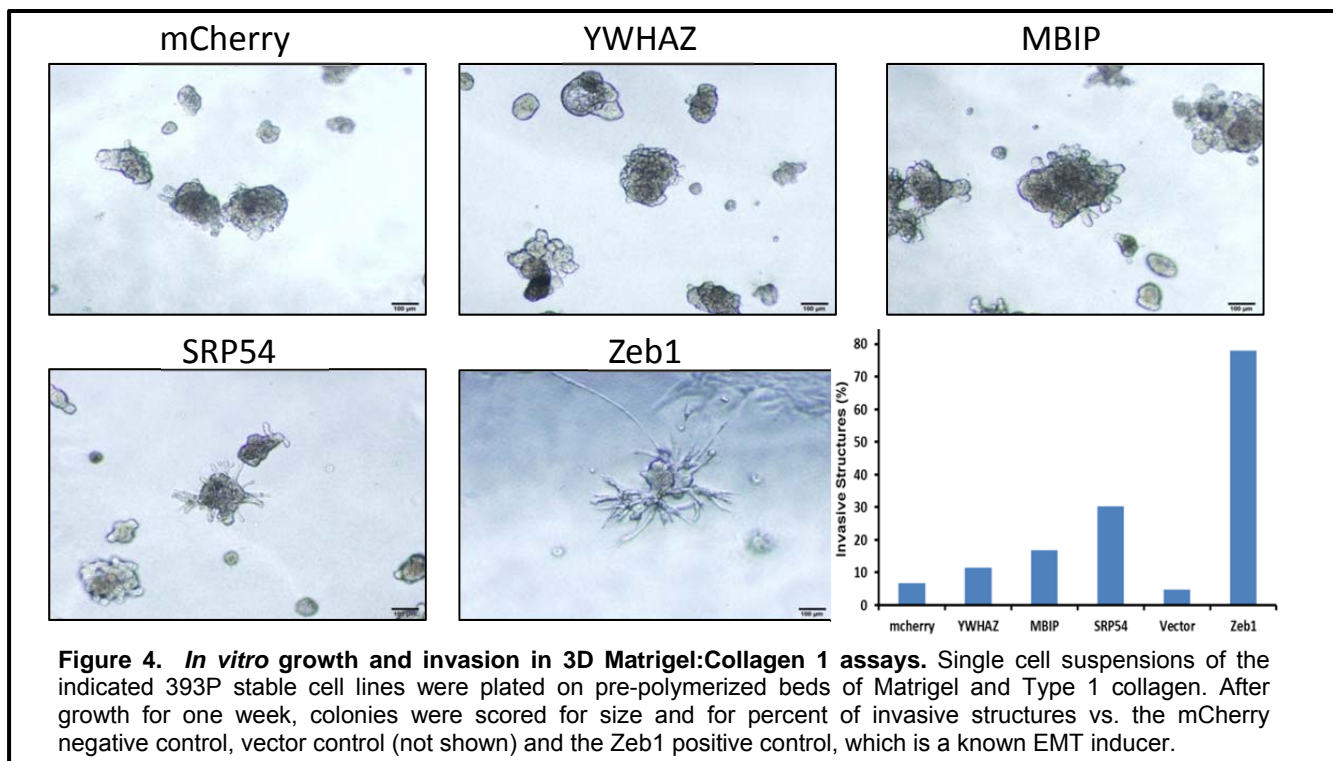
We initially proposed to complete the *in vitro* assays during the first year; however, we determined that screening would be more efficient if conducted in a “rolling” fashion whereby individual candidate pools (e.g., 20-30 gene sets of the 279 library panel) were entered into both the *in vitro* and *in vivo* screens simultaneously, thus eliminating the costly need to construct the transduced cell line panel on two separate occasions. Screening a subset of the 279 candidates has already revealed multiple genes with robust pro-invasion activity (Fig. 2B), and we are actively screening all 279 candidates and expect to complete the primary screens within the next 3 months (~01/01/2014).



Among the top scoring gene candidates screened thus far are *YWHAZ*, *MBIP* and *SRP54*, which were all subsequently validated by performing invasion and migration assays in the standard 24-well Boyden chamber system. Consistent with the primary screen in 96-well format, we observed significant increases in both migration and invasion for the 393P cells expressing each of the three genes compared to the mCherry negative control cells (Fig. 3A, B).



**Anoikis Resistance:** We have made great efforts to develop the proposed anoikis-resistance assays; however, our efforts have thus far been unsuccessful due to the fact that all cell lines assayed thus far (e.g., 393P, 393LN, etc) already exhibit robust anoikis resistance and are therefore not suitable as screening models. We are currently examining our newly-derived HBEC cell model described below to assess its performance in this assay, and we are also currently evaluating use of rat intestinal epithelial (RIE) cells, which is a screenable [3] “generalized” line highly sensitive to anoikis as demonstrated by RIE cells’ inability to proliferate under low cell attachment conditions. We anticipate determining whether or not anoikis screens will be feasible using these new cell systems within 2 months (~12/01/2013). Given this delay, we chose to enter cells stably transduced with ORF library virus into **3D culture assays** whereby cells are grown on a bed of extracellular matrix (Fig. 4). We have previously published that this method better mimics the *in vivo* condition by modeling the cell-matrix interactions, is scorable for cell growth and invasion, and can be modified by adjusting the composition and biophysical properties of the matrix [4-6]. As such we feel the 3D matrix assay is more robust and informative than the anoikis resistance assay, though we are continuing to develop the anoikis screens as described above. Our goal is to enter all ORF candidates into the 3D matrix assay given its tractability and robust output.

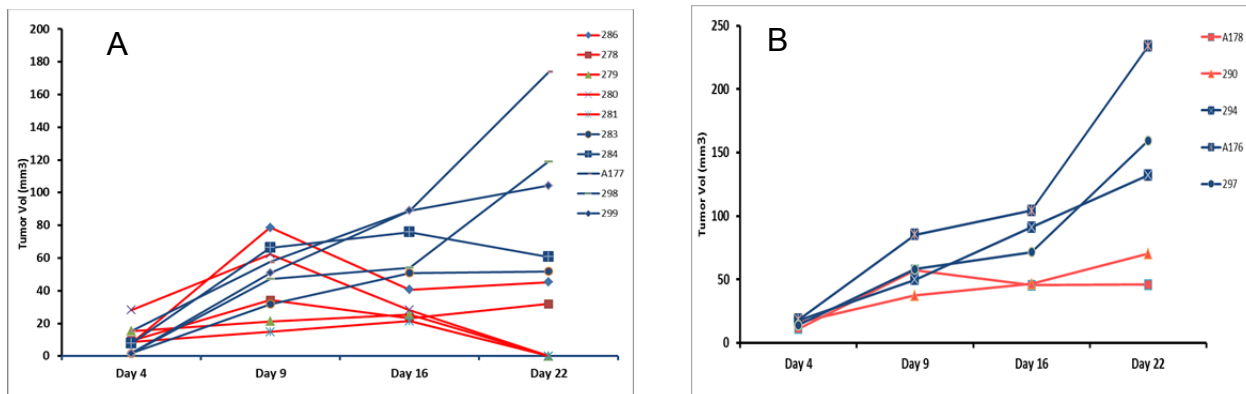


## Subaim 2.2. *In vivo* positive selection screens (proposed completion months 3-15).

In this Subaim we propose to perform *in vivo* metastasis screens with pooled viral-infected cells to positively select for single and combinatorial drivers of metastasis. Pilot assays demonstrated that subcutaneous injection of  $10^6$  cells provided an optimal period of primary tumor growth that could be monitored over ~6-8 weeks before the animals required sacrifice due to primary tumor burden (data not shown), thus allowing sufficient time for metastasis and detection of metastasis drivers. Although the cells can be inoculated orthotopically, the window of tumor growth is reduced to only a few weeks because the animals become tachypneic from the tumor burden in the lung, requiring earlier sacrifice. Since the goal of this *in vivo* screen is to recover distant metastases, the longer period of primary tumor growth is favored. After regulatory approval of the animal protocols (both the subcutaneous and orthotopic routes of injection are approved), cells infected and selected for stable expression of the ORFs in our library (as outlined above) were injected into the flank of syngeneic immunocompetent



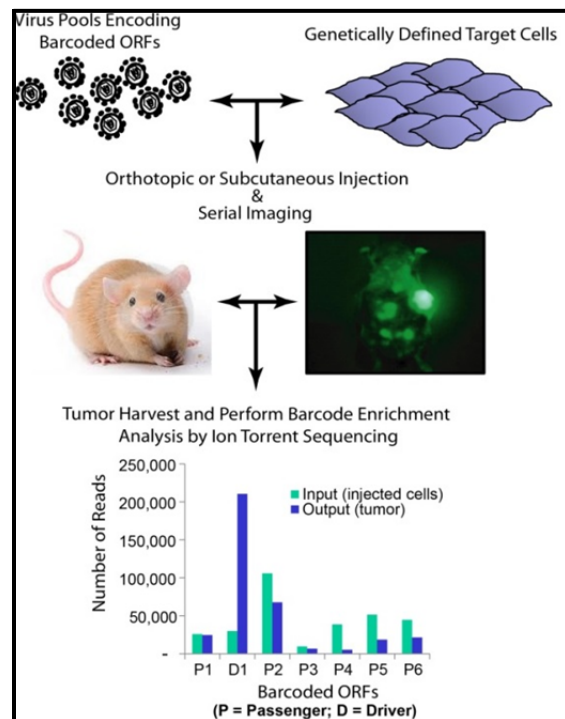
129Sv mice. Based upon pilot experiments of optimal tumor cell inoculum, each group of injections included 10 mice that were injected with a pool of  $10^6$  cells each, comprised of ORF expressing cell lines (50,000 cells/ORF) plus one additional mCherry control cell line used as an internal control. With each experimental pool we also injected 5 additional mice as negative control with  $10^6$  cells expressing mCherry only. Mice are observed for general health and the growth of primary subcutaneous tumors recorded (representative cohort shown in Fig. 4A,B). The animals are sacrificed when the tumor burden is excessive or the mice show apparent deterioration in general health.



**Figure 5. Subcutaneous primary tumor growth in experimental pool 1 and control mice. (A)** Mice were injected with 1 million cells comprised of 393P stable cell lines each expressing an individual ORF. Tumor volume was measured weekly for each mouse injected with the experimental pool. Each line plot represents one animal. **(B)** Mice were injected with 1 million cells expressing mCherry as control. Tumor volume was measured weekly for each mouse and graphed with each line plot representing one animal.

As discussed above, we have opted for a “rolling” screening process whereby individual candidate pools of the 279 library panel are being entered into both the *in vitro* and *in vivo* screens simultaneously to eliminate added costs and increase efficiency. We have entered approximately 75 of the 279 candidates into *in vivo* screens across multiple cohorts that are still under experimentation (i.e., tumors are actively growing and metastases pending). We are plotting survival and tumor growth curves (representative tumors shown in Fig. 4), and we will perform detailed animal necropsies, tissue collection and full pathological examination of all tissues in these animals upon their sacrifice.

Driver pro-viral inserts will be identified by their associated barcodes. As described above, we have made significant improvements to the size, workflow and efficiency of our clone processing platform over the duration of this study. In addition to our increased ability to generate mutant clones, one additional improvement afforded by this project relates to our molecular barcode detection strategy. We initially proposed to use Luminex beads to detect the barcodes incorporated into the vectors, but have switched to a sequencing platform because of improved quantification and decreased cost. This method allows us to identify driver genes from



**Figure 6. Enrichment analysis for drivers.** ORFs are introduced into cells (input) for injection into animals. Primary tumors and metastases (output) are harvested, macrodissected for identification of the proviral insert (by barcodes). Histogram serves as data example for positive enrichment analysis of driver (D; *HoxA1*) over passengers (P).



primary tumors and metastases using a common set of PCR primers that amplify barcodes from genomic DNA in parental (i.e., injected) cells, primary tumors and metastases. Barcode products are sequenced and quantitated by an Ion Torrent Personal Genome Machine (PGM), which can multiplex up to 96 tumor samples, thus significantly reducing overall cost. Enrichment is defined as candidates that are significantly higher in output (tumor or metastases) than input (injected cells). We observe that, relative to injected cells, tumors and metastases **positively select** driver genes and lose those with no role in tumor growth or metastasis (i.e., passengers). This technology has been fully optimized and was used to successfully identify drivers among non-pathogenic “passenger” genes in proof-of-concept studies supported by this project (Fig.6).

### Specific Aim 3: Clinicopathological prioritization and validation of top candidates.

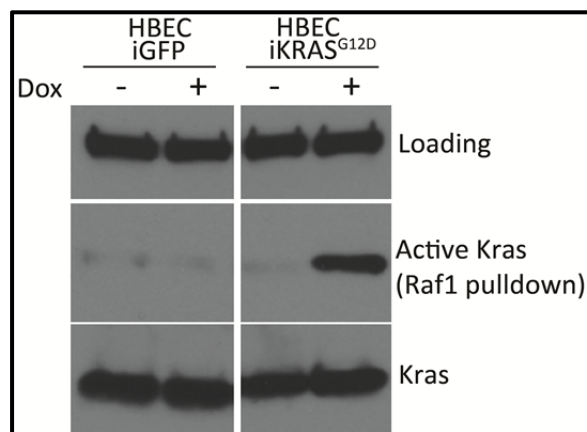
In this Aim we propose to take forward identified drivers from the *in vitro* and *in vivo* screens for clinical validation and prioritization by analysis of their expression in publically available datasets (e.g., TCGA and others, as well as in-house datasets with clinical annotation). We also propose to begin functional and mechanistic studies of clinically validated high priority candidates.

#### Subaim 3.1. Clinical validation and prioritization of metastasis genes (proposed completion during Year 2).

This work is still pending the results of screening, but will proceed within the next 3-6 months.

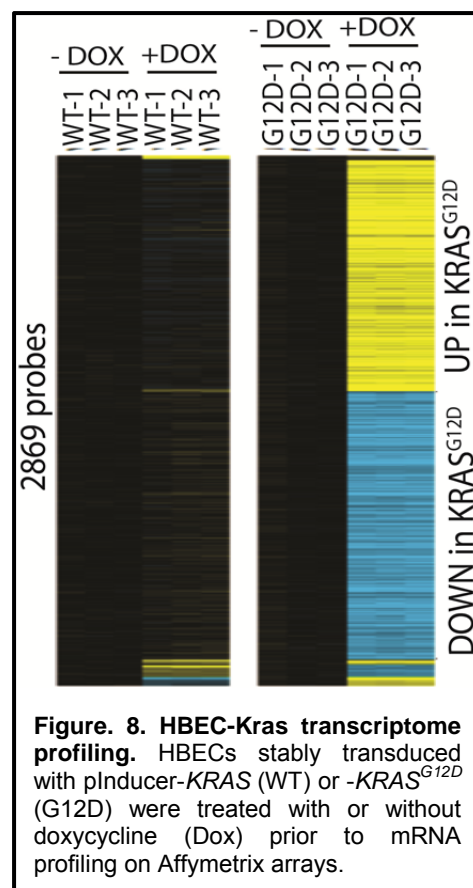
#### Subaim 3.2. Functional and mechanistic study of the lead metastasis genes (proposed completion during Year 2).

For oncogenic and metastasis driver validation assays, we proposed to use cancer cell lines for appropriate expression or knock-down studies. Depending upon the particular genes identified in the screens and the observed phenotypes, we will use human NSCLC cancer cell lines or other cancer cell types as appropriate. In preparation for these mechanistic studies we have also developed a unique normal human bronchial epithelial cell (HBEC) line that will be particularly useful. These HBECs provide a lineage-specific cell line engineered with signature aberrations to dissect driver mechanism in the proper genetic context. Parental HBECs, which were obtained from Dr. John Minna (U.T. Southwestern), were immortalized by expression of hTERT and contain aberrant Cdk4 and knockdown of p53. We used this line to develop a doxycycline inducible *KRAS*<sup>G12D</sup> construct (pInducer-*KRAS*<sup>G12D</sup>) based on our modified version of the pInducer vector system [7]. To test this cell model, virus carrying inducible GFP or oncogenic *KRAS*<sup>G12D</sup> was used to infect and select for stable HBEC cells, in which the expression of *KRAS* is controlled by treatment with or without doxycycline (Dox; 100ng/ml) for 2 days. Immunoblotting for *KRAS* activation by RAF pull down assays indicated significant *KRAS* activity in Dox-treated *KRAS*<sup>G12D</sup> cells, whereas no activity was detected in the absence of Dox and in GFP cells (with or without Dox; Fig. 7). It is noteworthy that, compared to GFP and –DOX controls, Dox-induced HBEC-i*KRAS*<sup>G12D</sup> cell lysates did not reveal significantly-elevated levels of total Kras protein owing to the low and titratable expression of the pInducer system. This suggests that our inducible model permits *KRAS* expression at a range near cell physiological levels that are desired for *KRAS* studies.



**Figure. 7. Immunoblot validation of HBEC-Kras model.** Immunoblot analysis of whole cell lysates extracted from HBEC cells engineered to express low levels of GFP or *KRAS*<sup>G12D</sup> afforded by used of the doxycycline (Dox) inducible system. Assessment Ras activation performed by RAF pull down assay.

In work funded through other sources, we have characterized these  $KRAS^{G12D}$  inducible HBEC lines. We performed mRNA transcription profiling for the WT  $KRAS$  and  $KRAS^{G12D}$  with or without doxycycline induction (Fig. 8). We found widespread gene expression changes upon  $KRAS^{G12D}$  induction, but not in the WT  $KRAS$  line. A comparison with a published  $KRAS$  associated gene expression signature [8] confirmed a significant enrichment for genes expressed in our  $KRAS^{G12D}$  lines ( $p = 2.1E-11$ ) upon doxycycline treatment. This lineage-specific immortalized human bronchial epithelial cell line with isogenic expression of mutant or wild-type  $KRAS$  therefore provides another important tool to test the phenotypic role of specific genes identified from the *in vitro* and *in vivo* screens, with the specific ability to determine the dependence of those genes on oncogenic  $KRAS$ .



**Figure 8. HBEC-Kras transcriptome profiling.** HBECs stably transduced with pInducer-KRAS (WT) or - $KRAS^{G12D}$  (G12D) were treated with or without doxycycline (Dox) prior to mRNA profiling on Affymetrix arrays.

## KEY RESEARCH ACCOMPLISHMENTS

- **ORF collection:** We expanded our screen platform to ~32,000 open reading frame (ORF; i.e., genes) clones. All ORFs are present in “entry” vectors compatible with Gateway-mediated recombineering, which allows high-throughput, robotics-driven ORF transfer to “destination” vectors for their expression in mammalian cells. Sequencing of all ORFs has enabled complete gene annotation, and our quality control criteria mandates 100% amino acid concordance with NCBI Reference entries.
- **High-throughput mutagenesis and molecular barcoding (HiTMMoB):** Enables high-throughput mutagenesis and molecular barcoding of our extensive ORF collection.
- **Development of a high-throughput *in vitro* invasion screen with dynamic range:** Facilitates rapid and quantifiable cellular phenotyping of the genes in the screen.
- **Robust 3D invasion assay:** Better mimics the 3D nature and extracellular matrix components found *in vivo*, while allowing manipulation of the matrix and easy scoring of the cellular phenotype.
- **Positive *in vivo* growth and metastasis screen:** Ability to directly test the *in vivo* role of genes on primary tumor growth and metastasis in a medium-throughput, pooled fashion due to the innovative barcoding techniques, use of sensitive sequencing and the positive selection of the *in vivo* screen.

## **REPORTABLE OUTCOMES**

- **pLentiEF6.3-puro:** We created a specialized lentiviral destination vector for the proposed studies. This vector permits accepts human wild-type and mutant ORFs via high-throughput recombination in the presence of DNA barcodes. The vector has an EF1a promoter compatible with expression in mouse and human cells, and also contains a puromycin selection cassette necessary for the cell lines used in this study.
- **Pro-invasion genes:** Primary screens and secondary validation have revealed a number of genes that robustly drive cell invasion (e.g., *YWHAZ*, *MBIP* and *SRP54*). A full list of functionally annotated scoring genes among the 260 candidates screened will be provided in our next report and resulting publications.
- **Mutant clone repository:** Cataloging of all sequenced-verified aberration clones constructed by this project which will be deposited at the DNASU Plasmid Repository (Arizona State University) to facilitate their distribution to the community thus maximizing their use.
- **Generation of an isogenic HBEC cell line with inducible mutant *KRAS* expression:** This cell system will allow us to test the oncogenic potential of the positive genes from the screen and the importance of the mutant *KRAS* background for their function.

## **CONCLUSIONS**

Our overall goal with this project is to establish a pipeline of robust screening techniques to functionally prioritize the data emerging from large-scale genomics efforts in lung cancer. Using a combination of *in vitro* and *in vivo* screens we will be able to identify and validate oncogene and metastasis drivers, explore the mechanistic basis for their function, and generate the pre-clinical cell and animal models needed for therapeutic targeting studies. As outlined in the Body of this report, we have successfully constructed the necessary libraries for this work, implemented both the *in vitro* and *in vivo* screens, and are putting into place the necessary tools for mechanistic studies of the lead candidate genes. We have made several technical improvements to our overall work plan and view this first year as a robust start to further screening of the TCGA data (which has only recently been released) that will extend past the life of this particular grant. We also feel that these techniques can be broadly implemented for functionalization of genomic data for other tumor types, e.g. pancreatic cancer.

## **REFERENCES**

1. Scott, K.L., et al., *Proinvasion metastasis drivers in early-stage melanoma are oncogenes*. Cancer Cell, 2011. **20**(1): p. 92-103.
2. Wardwell-Ozgo, J., et al., *HOXA1 drives melanoma tumor growth and metastasis and elicits an invasion gene expression signature that prognosticates clinical outcome*. . Oncogene, 2013. **Epup ahead of print**.
3. Douma, S., et al., *Suppression of anoikis and induction of metastasis by the neurotrophic receptor TrkB*. Nature, 2004. **430**(7003): p. 1034-9.
4. Gibbons, D.L., et al., *Contextual extracellular cues promote tumor cell EMT and metastasis by regulating miR-200 family expression*. Genes Dev, 2009. **23**(18): p. 2140-51.
5. Ahn, Y.H., et al., *ZEB1 drives prometastatic actin cytoskeletal remodeling by downregulating miR-34a expression*. J Clin Invest, 2012. **122**(9): p. 3170-83.
6. Gill, B.J., et al., *A synthetic matrix with independently tunable biochemistry and mechanical properties to study epithelial morphogenesis and EMT in a lung adenocarcinoma model*. Cancer Res, 2012. **72**(22): p. 6013-23.

7. Meerbrey, K.L., et al., *The pINDUCER lentiviral toolkit for inducible RNA interference in vitro and in vivo*. Proc Natl Acad Sci U S A, 2011. **108**(9): p. 3665-70.
8. Singh, A., et al., *A gene expression signature associated with "K-Ras addiction" reveals regulators of EMT and tumor cell survival*. Cancer Cell, 2009. **15**(6): p. 489-500.

## **APPENDICES**

None